

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2002, 12:34:11 ; Search time 2954.29 Seconds  
(without alignments)  
13199.933 Million cell updates/sec

Title: US-09-357-273A-1

Perfect score: 3629  
Sequence: 1 ccgctcgcagcgcgcgtca.....caggagccagcccgaaatc 3629

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues  
Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*

- 1: em\_estfun:\*
- 2: em\_esthum:\*
- 3: em\_estin:\*
- 4: em\_estom:\*
- 5: em\_estpl:\*
- 6: em\_estba:\*
- 7: em\_estro:\*
- 8: em\_estov:\*
- 9: em\_hic:\*
- 10: qb\_est1:\*
- 11: qb\_est2:\*
- 12: qb\_hic:\*
- 13: qb\_gss:\*
- 14: em\_gss\_fun:\*
- 15: em\_gss\_hum:\*
- 16: em\_gss\_inv:\*
- 17: em\_gss\_pln:\*
- 18: em\_gss\_pro:\*
- 19: em\_gss\_rnd:\*
- 20: em\_gss\_vrt:\*
- 21: em\_gss\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1019.4	28.1	1958	12 AK018505	AK018505 Mus muscu
2	554.2	15.3	680	11 BG080244	BG080244 H3051P08-
3	499.8	13.8	649	11 BE910769	BE910769 601661819
4	454.8	12.5	823	11 BG083506	BG083506 H3089H09-
5	447.4	12.3	482	11 BF846361	BF846361 PM1-EN006
6	431.2	11.9	638	11 BG711744	BG711744 pgl1n.pk0
7	333.4	9.2	380	11 BF846446	BF846446 PM1-EN006
8	293.6	8.1	920	10 BE736116	BE736116 601305874
9	277	7.6	279	11 BF868869	BF868869 I15-ET011
10	262.6	7.2	544	11 BG070677	BG070677 H3089H09-
11	254	7.0	587	11 BF582789	BF582789 602100922
12	241	6.6	556	10 AW642631	AW642631 cm20g11.w

13	227.8	6.3	449	10 AA413881	AA413881 vc53d02.s
14	227.2	6.3	512	10 A1733735	A1733735 z189c06.y
15	227.2	6.3	539	10 A1733738	A1733738 z191b12.y
16	216.2	6.0	565	10 AM809457	AM809457 MR4-ST012
17	212.4	5.9	677	11 BF964654	BF964654 602267732
18	200	5.5	342	10 AA794000	AA794000 vt40h10.s
19	189.6	5.2	740	11 BF582794	BF582794 602100927
20	189.2	5.2	320	13 A2600154	A2600154 1M0416B22
21	180.6	5.0	476	10 AA088547	AA088547 z189c06.r
22	178.6	4.9	539	11 B1394483	B1394483 pp1n.pk0
23	157.2	4.2	421	10 A1734961	A1734961 rE05c11.x
24	151.2	4.2	422	10 A1734969	A1734969 at05d10.x
25	139.4	3.8	348	10 AM846947	AM846947 RC3-CT019
26	137.6	3.8	426	10 AU016754	AU016754 AU016754
27	136.2	3.8	775	13 CNS02PEK	AL207893 Tetradon
28	132.4	3.6	170	11 BF817483	BF817483 RC5-CT014
29	114.6	3.2	403	10 AA619271	AA619271 v050h02.r
30	114.6	3.2	531	13 A2294936	A2294936 RPT-23-9
31	114.6	3.2	1101	13 CNS04008	AL308249 Tetradon
32	114	3.1	561	11 BG710600	BG710600 pgl1n.pk0
33	104.2	2.9	133	11 BF995998	BF995998 MR1-GN017
34	104	2.9	780	11 BG784666	BG784666 SEAMC004
35	100.4	2.8	571	13 A2626875	A2626875 1M0467D19
36	99	2.7	482	13 A2080607	A2080607 RPT-23-3
37	98	2.7	948	13 CNS0011B	AL074608 Drosoph11
38	93.6	2.6	322	10 A1986903	A1986903 rs19a12.y
39	92.8	2.6	561	13 A0738301	A0738301 HS-2260.B
40	92	2.5	999	13 CNS06MSD	AL405875 T3 end of
41	86	2.4	1134	13 CNS02345	AL179006 Tetradon
42	85	2.3	348	13 B51730	B51730 C1T978SK-A-
43	82.4	2.3	339	10 AA755189	AA755189 vq62b05.r
44	82.2	2.3	283	10 AA102368	AA102368 z191b12.r
45	81.2	2.2	721	13 A2711940	A2711940 RPT-24-8

#### ALIGNMENTS

RESULT 1	AK018505	1958 bp	mRNA	HTC	05-JUL-2001
LOCUS	AK018505				
DEFINITION	Mus musculus adult male colon cDNA, RIKEN full-length enriched library, clone:9030414B18, full insert sequence.				
ACCESSION	AK018505.1 GI:12858236				
VERSION	1				
KEYWORDS	CAP trapper.				
SOURCE	Mus musculus (strain:C57BL/6J) adult male colon cDNA to mRNA, clone:lib:RIKEN full-length enriched mouse cDNA library				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	1 (bases 1 to 1958)				
TITLE	Carninci, P. and Hayashizaki, Y.				
JOURNAL	High-efficiency full-length cDNA cloning				
MEDLINE	Methods in enzymology. 303, 19-44 (1999)				
PUBMED	99279253				
REFERENCE	2 (bases 1 to 1958)				
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to				
JOURNAL	prepare full-length cDNA libraries for rapid discovery of new genes				
MEDLINE	Genome research. 10 (10), 1617-1630 (2000)				
PUBMED	20499374				
REFERENCE	3 (bases 1 to 1958)				
AUTHORS	Kono, H., Akiyama, J., Nishi, K., Kitunai, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishigaki, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Iwawa, M., Ohara, E., Matshiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,				



Db	1091	TCGAGCGACGACCTCTTCCTTGGTGGAAAGGCCCCAGACAGATGGCGTCCACCATTTGAGG	1150
Qy	1080	caagagggaatctgtatataacggccagcgagcgtcaagttagtcccgactccaag	1139
Db	1151	CMAAGGAGAGTGTGTATCTACTCCAGCACAAGCTTCAAGTTTGACCTGGACTCAAAAGG	1210
Qy	1140	caagaacaagctcaactacttgaagaaattacttgcgtctctgatalagaaacacatlgaaacccc	1199
Db	1211	GAACAGCGAACCTAATCTACTTGGAGATTAATCTGGCTTCTCATTAGGACACATATGAACATCC	1270
Qy	1200	actgtcgcgtctacaagaatgcttgaagaattcccaacaatctcaaccaacatcggga	1259
Db	1271	TCGTGTGCATCCACCAAGATGCTGGAGAGATTTCCTAACAACCTGTGCCCAAAACATCGAGA	1330
Qy	1260	aaatgtatctccgtcgtgattcagagaanaaagcctttaggaagtta	1306
Db	1331	AAATGTATTTCTCGCTGATTCACAAAAAGAGAGCTTTTGAGGAAGTGA	1377

SUITE		2
LOCUS	BG080244	
DEFINITION	680 bp mRNA EST	26-JAN-2001
ACCESSION	H305ID08-5 NIA Mouse 15K cDNA Clone Set Mus musculus CDNA clone H305ID08 5', mRNA sequence.	
VERSION	BG080244	
KEYWORDS	EST.	
SOURCE	house mouse.	
ORGANISM	Mus musculus	
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
TITLE	Kargul,G.J., Dudekula,D.B., Qian,Y., Lim,M.K., Jaraedt,S.A., Tanakaka,T.S., Carter,M.G. and Ko,M.S.H.	
JOURNAL	Verification and Initial annotation of NIA mouse 15K cDNA clone set unpublished (2001)	
COMMENT	Other ESTs: H305ID08-3 Contact: George J. Kargul Laboratory of Genetics National Institute on Aging/National Institutes of Health 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA Email: cdna@igsun.grc.nia.nih.gov This clone set has been freely distributed to the community. Please visit http://IGsun.grc.nia.nih.gov/cDNA/15k.html for details. Plate: H305I row: D column: 08 Seq primer:-21M13 Reverse High quality sequence stop: 680  POLYA-No.	
FEATURES	location/Qualifiers	
SOURCE	l..680 <code>/organism= "Mus musculus"</code> <code>/strain= "C57BL/6J"</code> <code>/db_xref= "taxon:10090"</code> <code>/clone= "H305ID08"</code> <code>/clone_1lb= "NIA Mouse 15K cDNA Clone Set"</code> <code>/sex= "Clones arrayed from a variety of cDNA libraries"</code> <code>/dev_stage= "Clones arrayed from a variety of cDNA libraries"</code> <code>/lab_host= "DH10B"</code> <code>/note= "Vector: pSPORT1; Site_1: SalI; Site_2: NotI. This clone is among a rearranged set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos , and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source cDNAs are cloned unidirectionally with Oligo(dT )-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, Proc.Natl. Acad. Sci. U S A, 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression</code>	

BASE COUNT 175 a 192 c 142 t 1 others  
ORIGIN 1

Query Match	Similarity	15.3%	Score 554.2	DB 11	Length 680
Best Local	Similarity	89.5%	Pred. No. 1.7e-106		
Matches	607	Conservative	0	Mismatches	69
				Indels	2
				Gaps	1
QY 901	ccgttccccaaagagacagagagccaaagagcaagctgacgcgccactgtatgttggaaa				960
Db 2	ccatttccccaaagagacagagagccaaagacaaagctaacgcttactctgtatgttggaa				61
QY 961	tacttaccagccttatagtccttccctcaatgtgtacacagaggggtgtctgtgtccc				1020
Db 62	tatttcaccagccttatagtccttccctcaatgtgtacacagaggggtgtctgtgtccc				121
QY 1021	cgcgcagcagcaccttcccttctgttggaaagggcccccagacgtagtgggttcaacctcggggac				1080
Db 122	cgagcgacgactcttcttcttggaaagggcccccagacgtagtgggttcaacctcggggac				181
QY 1081	aagggggaggtgtgtatcacgcgccaaagcagagacgctcaaatgttatcccgagactaaaagc				1140
Db 182	aaagagagaggtgtgtatcacgcgccaaagcagagacgctcaaatgttatcccgagactaaaagc				241
QY 1141	aagaaacagctcaactacttgaagaattactgtgtctgtatagagacacatgaaaccca				1200
Db 242	aagacacagctcaactacttgaagaattactgtgtctgtatagagacacatgaaaccca				301
QY 1201	ctgtctgtcgttaccagaatgtcttggagagatttcccaacaactaccaccaatctggagaa				1260
Db 302	ctgttggatctccaccaaagatgtcttggagagatttcccaacaactaccaccaatctggagaa				361
QY 1261	aatgtgatctccgtctgattcagagaaaagagcttggagaaattcaacctggtgtgac				1320
Db 362	aatgtgatctccgtctgattcagagaaaagagcttggagaaattcaacctggtgtgac				421
QY 1321	cagacttcaagaaaagcacccttaaccaacgtgtcttgcggagatgttggagagaagcccgccat				1380
Db 422	cagacttcaagaaaagcacccttaaccaacgtgtcttgcggagatgttggagagaagcccgccat				481
QY 1381	gcccttccgcggcccgagagggcccccgttggactccaigtctaaggaactgtctacacatc				1440
Db 482	gcccttccgcggcccgagagggcccccgttggactccaigtctaaggaactgtctacacatc				541
QY 1441	ctgaagcacccttccgtctgtatgtgtggg--tggacctcaatcaacttccccctagaca				1498
Db 542	ctgaagcacccttccgtctgtatgtgtggg--tggacctcaatcaacttccccctagaca				601
QY 1499	tgcattcagcagcagcagcttccagcaaccaagcaatgtccagaagaacttgaagaatccac				1558
Db 602	tgcattcagcagcagcagcttccagcaaccaagcaatgtccagaagaacttgaagaatccac				661
QY 1559	tccctcagcagcagcagcagc 1576				
Db 662	tccctcagcagcagcagcagc 679				
RESULT 3					
LOCUS	BE910769	649 bp	mRNA	EST	29-SEP-2000
DEFINITION	601651819p1 NCI_CGAP_Mam1 Mus musculus CDNA clone IMAGE:3962092	5'			
ACCESSION	BE910769				
VERSION	BE910769.1	GI:10407697			
KEYWORDS	EST.				
SOURCE	house mouse.				
ORGANISM	Mus musculus				

REFERENCE 1 (bases 1 to 649)  
 AUTHORS NIH-MGC <http://mgi.ncl.nih.gov/>  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: Gilbert Smith, Ph.D.  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Plate: L14M9128 row: e column: 05  
 High quality sequence stop: 648.

FEATURES  
 source  
 1..649  
 Location/Qualifiers

/organism="Mus musculus"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:3962092"  
 /clone\_lib="NCI-CCAP\_Mam1"  
 /tissue\_type="tumor, biopsy sample"  
 /dev\_stage="3 months, virgin"  
 /lab\_host="DH10B"  
 /note="Organ: mammary; Vector: PCMV-SPORTe; Site:1: Salt; Site:2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"  
 BASE COUNT 154 a 181 c 172 g 142 t  
 ORIGIN

Query Match 13.8%; Score 499.8; DB 11; Length 649;  
 Best Local Similarity 85.8%; Pred. No. 4.7e-95;  
 Matches 555; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

2270 tcaagccgcagatctgggagcctgagcagaaagcgcgagatcgcagatgtagcg 2329  
 1 ttcagccgcggttcagctgacgtgacgtgacgtgacgtgacgtgacgtgacgtg 60  
 2330 aagactgtatagaagaacccctaccacacggtgagacatcttctcgaagctgcgtcttt 2389  
 61 aagactgtatagaagaacccctaccacacggtgagacatcttctcgaagctgcgtcttt 120  
 2390 actaagtggtctctgagggcagaccccttttgcaagtccttcgagcgagcgcaaca 2449  
 121 actatgtcattctctgagggcagaccccttttgcaagtccttcgagcgagcgcaaca 180  
 2450 tccctcgggtgagcctgagcctgagcctgagcctgagcctgagcctgagcctgagc 2509  
 181 tccctcgggtgagcctgagcctgagcctgagcctgagcctgagcctgagcctgagc 240  
 2510 cagcgaatctgataagaagatgtagtgcagatgtagtgcagatgtagtgcagatgtag 2569  
 241 ctgcgtgaattgataagaagatgtagtgcagatgtagtgcagatgtagtgcagatgtag 300  
 2570 agtgtgtcaaacacccgtctcttctgagcctgagcctgagcctgagcctgagcctgag 2629  
 301 acgtgtgaaacacccgtctcttctgagcctgagcctgagcctgagcctgagcctgag 360  
 2630 tgaagcagcgaatagaaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaaag 2689  
 361 tgaagcagcgaatagaaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaaag 420  
 2690 ggcggagagcgcgtgtgtaagatgtagtgcagatgtagtgcagatgtagtgcagatgtag 2749  
 421 ggcggagagcgcgtgtgtaagatgtagtgcagatgtagtgcagatgtagtgcagatgtag 480  
 2750 accctgttaaatctcagagcctataaagtggtctctcagagatctcctccgagcattga 2809  
 1 ttt

Db 481 ATCTGGCAATTCAGAACCTACCAAGGTGCTCTGTGAGAGACCTCTCCAGCCATGA 540  
 Qy 2810 gaataagaagac 2869  
 Db 541 GAAACAGAAAC 600  
 Qy 2870 tccccgacgactctgtgtgctactacgtcgtctcctccccacacac 2916  
 Db 601 TCCCGATGACTTGTGGCTACTTATTCCTCCGTCCCACTCCAC 647

RESULT 4  
 BG083506 823 bp mRNA EST 26-JAN-2001  
 LOCUS H3089H09-5 NIA Mouse 15K CDNA Clone Set Mus musculus CDNA clone  
 DEFINITION H3089H09 5', mRNA sequence.  
 ACCESSION BG083506  
 VERSION BG083506.1 GI:12566162  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 823)  
 AUTHORS Kargul, G.J., Dudekula, D.B., Qian, Y., Lim, M.K., Jaradat, S.A., Tanaka, T.S., Carter, M.G., and Ko, M.S.H.  
 TITLE Verification and initial annotation of NIA mouse 15K CDNA clone set  
 JOURNAL Unpublished (2001)  
 COMMENT Other ESTs: H3089H09-3  
 Contact: George J. Kargul  
 Laboratory of Genetics  
 National Institute on Aging/National Institutes of Health  
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
 Email: [cdna@igsun.grc.nia.nih.gov](mailto:cdna@igsun.grc.nia.nih.gov)  
 This clone set has been freely distributed to the community. Please visit <http://igsun.grc.nia.nih.gov/cdna/15k.html> for details.  
 Plate: H3089 row: H column: 09  
 Seq primer: -21M13 Reverse  
 High quality sequence stop: 823  
 POLYA-No.

FEATURES  
 source  
 1..823  
 Location/Qualifiers

/organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="H3089H09"  
 /clone\_lib="NIA Mouse 15K CDNA Clone Set"  
 /sex="Clones arrayed from a variety of CDNA libraries"  
 /dev\_stage="Clones arrayed from a variety of CDNA libraries"  
 /lab\_host="DH10B"  
 /note="Vector: pSPORT1; Site:1: Salt; Site:2: NotI; This clone is among a rearranged set of 15,247 clones from 11 embryo CDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary CDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT) Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental CDNA microarray, 2000, Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2) Large-scale CDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."  
 BASE COUNT 223 a 204 c 191 g 187 t 18 others  
 ORIGIN

Query Match 12.5%; Score 454.8; DB 11; Length 823;  
Best Local Similarity 87.4%; Pred. No. 1.5e-85;  
Matches 498; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

```
QY 760 taagcctccctctgtgtgagccttattgtctgacagcgaggaaggtctgaggaagtatg 819
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 1 TATGCTCTCCTGCTGTGTGCTCTTCTACGTCGTGACAGGGGAGCTCTGAGAAAGTGCTG 60
QY 820 cccatcatgtctgtgtgagacccctgtgctatctgaccttattgtctgggaagtgtgag 879
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 61 CACATCAACGTTGCTGTAGAGACTTACGCTTCTGACCTTCTGAGCTCTGAGAGGGG 120
QY 880 cccatcaacaaggtggaagaccgtctccccaagagagacagagcccaagactgagc 939
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 121 CCCATCACCACCAATGGATTATTCATTCGCCAAGAGACAGAGGCCAAGAGCAAGCTAACG 180
QY 940 cccactctgtatgtctggaagaatactctaccagcctctatgctctccctcaatgtgtacac 999
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 181 CCTACTCTGTATGTTGGGAAAGTATTCACACGCTCTATGCTCTCCCTCAATGGTGACAT 240
QY 1000 gaggagggtgtctgtctgtcccgagcagacacttctctgtctggaagggcccaagact 1059
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 241 GAGGGGGTGTGCTGTGCTCTGAGGACACACTCTTCTTCTGGAAGGCCCCAGACACA 300
QY 1060 gatgagctacacatcgaggagagagggaggtgtgtatcaagcccaagcagagctcaag 1119
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 301 GATGGCGCTACCATTTGGAGACAAAGAGAGGTGTGTATCTCCACAGCAAGACCTCAAG 360
QY 1120 ttgtatcccgagctcaaaagaagaagaacgaactcaactcttgaggaaatctactgtctctg 1179
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 361 TTTGACCCCTGAGCTCAAAAGGAGAGCAAGCTGAACTACTTGAAGAAATTAAGTGCTTCTC 420
QY 1180 ataggaacacatgaagaacccactgtctgtcttaccagaatgtctggagagatttcccaac 1239
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 421 ATAGAGACACCAATGAACCTCTGCTGTGATCCACCAAGATCTGTGAGAGATTTTCTTAAC 480
QY 1240 aatctcaacaacatcgaggaagaatgtgtatctctgtctgtatctgagagaagaagctttgag 1299
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 481 AACCTGCCCAACATCGAAGAAATGTGATTCCTGCTGATTCAGAAAAGAGAGCTTTGAG 540
QY 1300 gaagtatcaacctgtgtgacagacttca 1329
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 541 GAAAGATCAAACTCAGATGTTCAGACTTGA 570
```

RESULT 5

LOCUS	BF846361	482 bp	mRNA	EST	16-JAN-2001
DEFINITION	PM1-EN0065-231000-002-c07	EN0065	Homo sapiens	CDNA, mRNA sequence.	
ACCESSION	BF846361				
VERSION	BF846361.1	GI:12233615			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
	1 (bases 1 to 482)				
	Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,				
	Nagal, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.R.,				
	Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,				
	Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare				
	, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and				
	Simpson, A.J.				
TITLE	Shotgun sequencing of the human transcriptome with ORF expressed				
JOURNAL	Sequence tags				
MEDLINE	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)				
COMMENT	20202663				
	Contact: Simpson A.J.G.				
	Laboratory of Cancer Genetics				
	Ludwig Institute for Cancer Research				
	Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,				
	Brazil				

Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?l1=PM1&l2=PM1-EN0065-  
231000-002-c07&t3=2000-10-23&t4=1)  
Seq primer: puc 18 forward  
High quality sequence stop: 480.

FEATURES

source

1..482

Location/Qualifiers

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_lib="EN0065"

/dev\_stage="Adult"

/note="Organ: lung\_normal; Vector: puc18; Site.1: Smat;  
Site.2: Smat; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the puc 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."

BASE COUNT 118 a 130 c 137 g 97 t

ORIGIN

Query Match 12.3%; Score 447.4; DB 11; Length 482;  
Best Local Similarity 97.5%; Pred. No. 4.8e-84;  
Matches 465; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

```
QY 2209 atcaagggcatatctctccactcttgccctctgcaagaagctgagtgagcagacagt 2268
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 7 ATCCAGGCCATATCTCCGACTTTGGCCCTCTGCAGAACTGGCAGTGGCAGACACACT 66
QY 2269 ttcagccgcagatctgagggtgctctgacacagaagctgagtcctccagagatctgagc 2328
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 67 TTCAGCCGCGATCTGGGGTGTGCTGCGACACAGAGCTGGATGCTCCAGATGCTGAGC 126
QY 2329 gaagactgtaaagagaacacctaccacagagtgagacatctttctgagagctgagcttt 2388
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 127 GAAGACTGTAAAGAGAACCTTACCTTACACGGTGAATTTTGTGACAGCTGCGCTTTT 186
QY 2389 taactcgtgtctctgagggagcagcacccttctgcaagctccctgagggagggccaac 2448
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 187 TACTAGCTAATCTCTGAGGGAGCCACCCTTTGGCAAGTCCCTGACAGGGAGGCCAAC 246
QY 2449 atcctcctggtgtgctctgagccttgactgtctgcaaccagagaagcaagaagtcatt 2508
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 247 ATCTCTCTGTGGTGGCTGACGCTTGACTGCTTG -ACCCAGAGAAACACAGAACGTCATTT 305
QY 2509 gacagagaattgatgagaagatlttgatgagatcctccagaagaagccctcagaagac 2568
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 306 GCACGTGAATTGATGAGAAAGATGATGGATGATCTTCAAGAAAGCCCTCACCGAAG 365
QY 2569 gacgtgtctcaaacaccgcgtctctctggaagcctagagaagcagctcagcttctcaggag 2628
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 366 CACGTGCTTAAACACCCGTTTCTTGAGACCTTAAGAGAACACACTCTCAGTGCCTTCAGAGAC 425
QY 2629 gtgagcgacagataaagaagaatccctgtgataggcccgatctgtgaagcagttagag 2685
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 426 GTGACGCGACAGAAATGAAAGAAATCCCGATGTGCGCATGTGAAGAGATGAAGAG 482
```

RESULT 6

LOCUS	BG711744	638 bp	mRNA	EST	08-MAY-2001
DEFINITION	pglin.pK009.e11	Normalized liver library	Gallus gallus	CDNA clone	
ACCESSION	BG711744				
VERSION	BG711744.1	GI:14005694			
KEYWORDS	EST.				
	pglin.pK009.e11	Similar to gi14557569 ref NP_001424.1	ER to		
		nucleus signalling 1 (Homo sapiens)	gbl1AAC25991.1 (AF059198)		
		protein kinase/endonuclease (Homo sapiens)	g, mRNA sequence.		



DB	LOCUS	REFERENCE	FEATURES
Db	146	GAAGACTGTGAAGGAGAACCCCTACCTACACAGGGTGGACACATCTTTTCTGACAGGCTGCTCTTT	205
Qy	2388	tactacgtgtgtctctgtgagggcagccacccttltggcaagtcctctgacgagcgagcccaac	2448
Db	206	TACTACGTGAATCTCTGAGAGGCGACCCACCCCTTTTGGCAAGTCCCTCAGCGAGCGCAAC	265
Qy	2449	atccctctggagtgctctgacgaccttactgcgtctgacccagagaagcacagaagtcactt	2508
Db	266	ATCCTCTCTGGGTCGCTGCAGCCTTACTGCTTGCCACCCAGAGAACACAGACACTCATTT	325
Qy	2509	gcacgagaa-ttgatagagaagaatgatgttcgcatgtgatcctcagaacgccccctca	2562
Db	326	GCACCTGAACCTGTGATAGAGAAAGATGATTGGCAGTGGCTCCTCAGAAACCCCTTCA	380
RESULT	8		
BE736116		BE736116	920 bp mRNA EST 15-SEP-2000
LOCUS		601305874P1 NIH_MGC_39	Homo sapiens cDNA clone IMAGE:3640266 5',
DEFINITION		mRNA sequence.	
ACCESSION		BE736116	
VERSION		BE736116.1	GI:10150108
KEYWORDS		EST.	
SOURCE		human.	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.	
TITLE		1 (bases 1 to 920)	
JOURNAL		NIH-MGC http://mgs.nci.nih.gov/.	
COMMENT		National Institutes of Health, Mammalian Gene Collection (MGC)	
		Unpublished (1999)	
		Contact: Robert Strausberg, Ph.D.	
		Email: cs9apbs-remail.nih.gov	
		Tissue Procurement: ATCC	
		cDNA Library Preparation: Ling Hong/Rubin Laboratory	
		cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)	
		DNA Sequencing by: Incyte Genomics, Inc.	
		Clone distribution: MGC clone distribution information can be	
		found through the I.M.A.G.E. Consortium/LNL at:	
		http://image.lnl.gov	
		Plate: LINC3144 row: c column: 19	
		High quality sequence stop: 657.	
FEATURES		location/qualifiers	
source		1..920	
		/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:3640266"	
		/clone_lib="NIH_MGC_39"	
		/tissue_type="adenocarcinoma"	
		/lab_host="DH10B (phage-resistant)"	
		/note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;	
		Site_2: EcoRI; cDNA made by oligo-dT priming.	
		Directionally cloned into EcoRI/XhoI sites using the	
		following 5' adaptor: GGCACGAG(G). library constructed	
		by Ling Hong in the laboratory of Gerald M. Rubin	
		(University of California, Berkeley) using ZAP-cDNA	
		synthesis kit (Stratagene) and Superscript II RT (Life	
		Technologies)."	
BASE COUNT		147 a 308 c 286 g 179 t	
ORIGIN			
Query Match		8.1%; Score 293.6; DB 10; Length 920;	
Best Local Similarity		65.3%; Pred. No. 1.4e-51;	
Matches	462; Conservative	0; Mismatches 244; Indels 2; Gaps 2;	
Qy	2097	cttgctgcaagcagcacacctcggcgctctggccacactccatccctcaaatcgttacag	2156
Db	8	CGTGTGCGAGCAGCTGATGCTGTGCGCTGGCCACCTGCACCTCTTAACATATAGTGACCG	67
Qy	2157	agacctaaagccacacaaatcctcatatcatctgtcccaatgcacagcggaatcaagc	2216
Db	68	GGACCTGAAGCAGGAATATTTCTATCATCCCGGGGCTGACACAGCGGCTTGGGAGACT	127

Db	Accession	Version	Keywords	Source	Organism
OY	2217				
Db	128				
OY	2277				
Db	188				
OY	2337				
Db	248				
OY	2397				
Db	308				
OY	2457				
Db	368				
OY	2517				
Db	428				
OY	2577				
Db	488				
OY	2637				
Db	548				
OY	2657				
Db	608				
OY	2757				
Db	667				
RESULT	9				
LOCUS	BF68869/c				
DEFINITION	115-ER0119-181000-175-cl1 ER0119				
ACCESSION	BF68869				
VERSION	BF68869.1				
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
REFERENCE	1 (bases 1 to 279)				
AUTHORS	Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,				
	Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,				
	Goldman,G.H., Carvalho,A.F., Metsukuna,A., Bais,G.S., Simpson,D.H.				
	Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare				
	,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and				
	Simpson,A.J.				
TITLE	Shotgun sequencing of the human transcriptome with ORF expressed				
	sequence tags				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)				
MEDLINE	2002063				
COMMENT	Contact: Simpson A.J.G.				
	Laboratory of Cancer Genetics				
	Ludwig Institute for Cancer Research				
	Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,				
	Brazil				
	Tel: +55-11-2704922				
	Fax: +55-11-2707001				
	Email: asimpson@ludwig.org.br				
	This sequence was derived from the FAPESP/LICR Human Cancer Genome				







